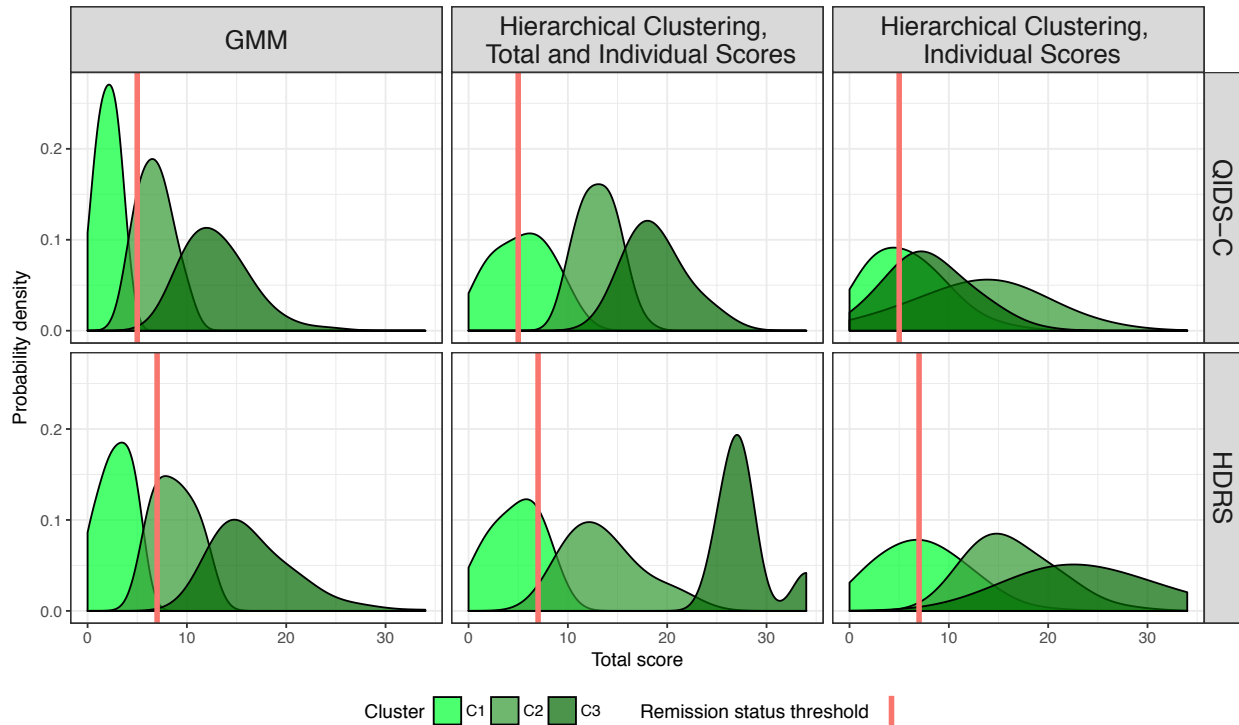# SUPPLEMENTARY FIGURES

Fig. S1: Comparison of depressive symptom clustering behavior, using various approaches. Shown here are the probability density functions (PDFs) of our clustering approach when univariate total depression severity scores at 8 weeks were used. Results with a Gaussian mixture model (GMM) are shown in the first column. The middle and third columns show clustering of patients done via hierarchical clustering approaches with multivariate data that comprise individual item responses at 8 weeks, both with (middle column) and without (third column) total depression severity scores at 8 weeks. The probability densities in each figure are represented on the y-axis, and depressive symptom scores derived using the QIDS-C (first row) and HDRS (second row) rating scales are represented on the x-axis. The threshold for remission for each depression rating scale is indicated by the red vertical line in each plot. The ecological validity of the GMM clustering approach using univariate depression severity scores is represented by the fact that the C1 clusters for QIDS-C and HDRS fall entirely within the range of scores that define remission. Neither of the two hierarchical clustering approaches yielded C1 clusters that fell entirely within the range of scores that define remission for either depression rating scale.
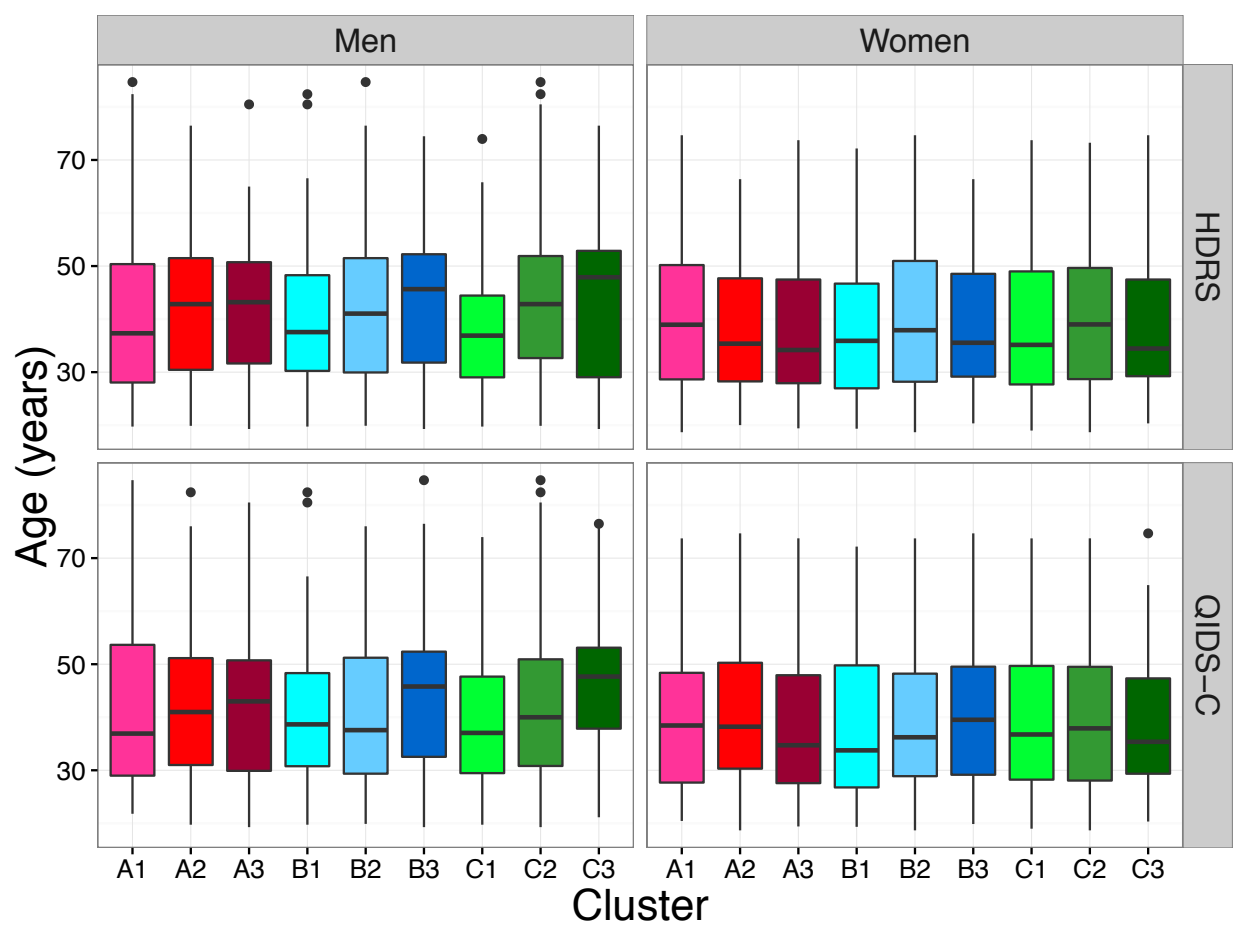
Fig. S2: Comparison of mean ages for men and women in clusters with comparable symptom severity at baseline, 4 weeks, and 8 weeks.
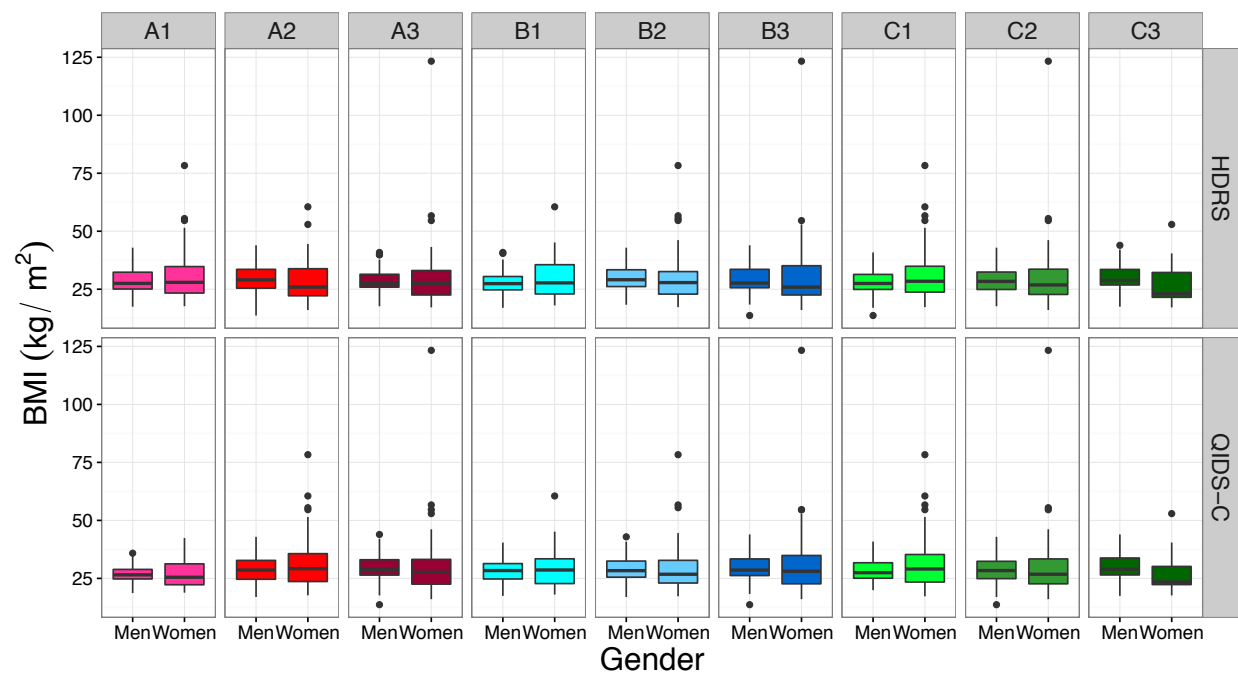
Fig. S3: Comparison of mean body mass indices (BMI, kg/m^2) for men and women in clusters with comparable symptom severity at baseline, 4 weeks, and 8 weeks.
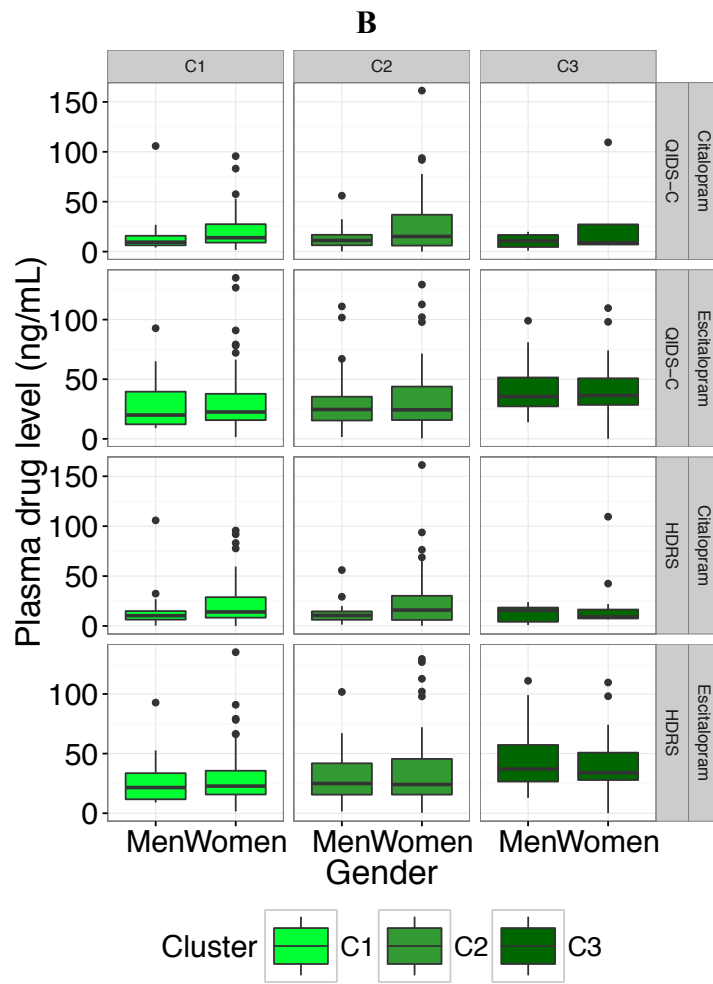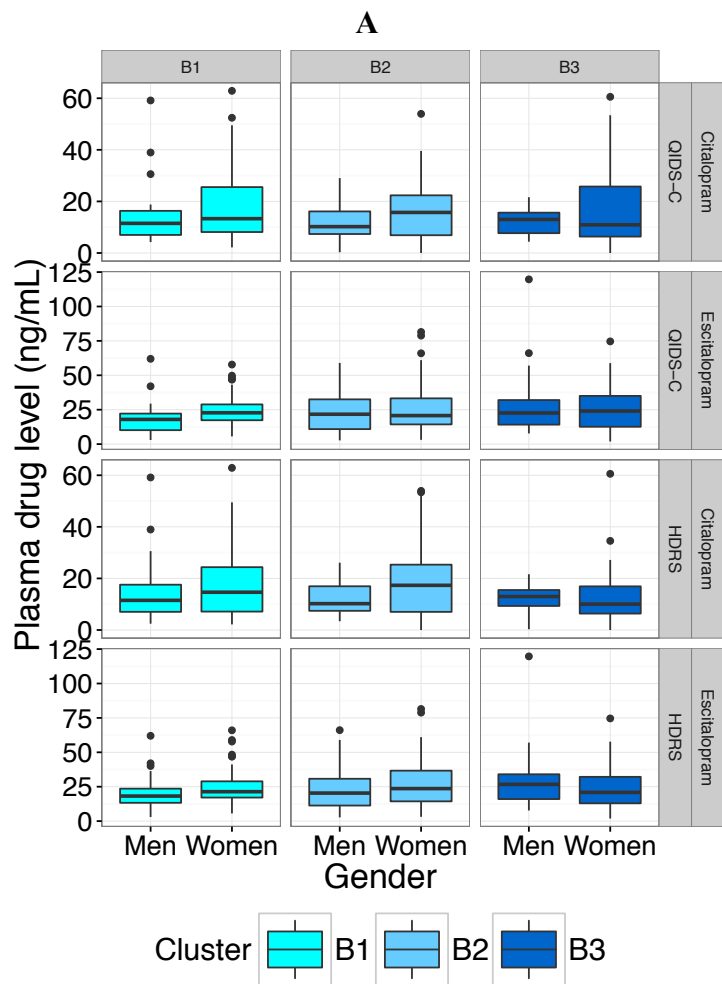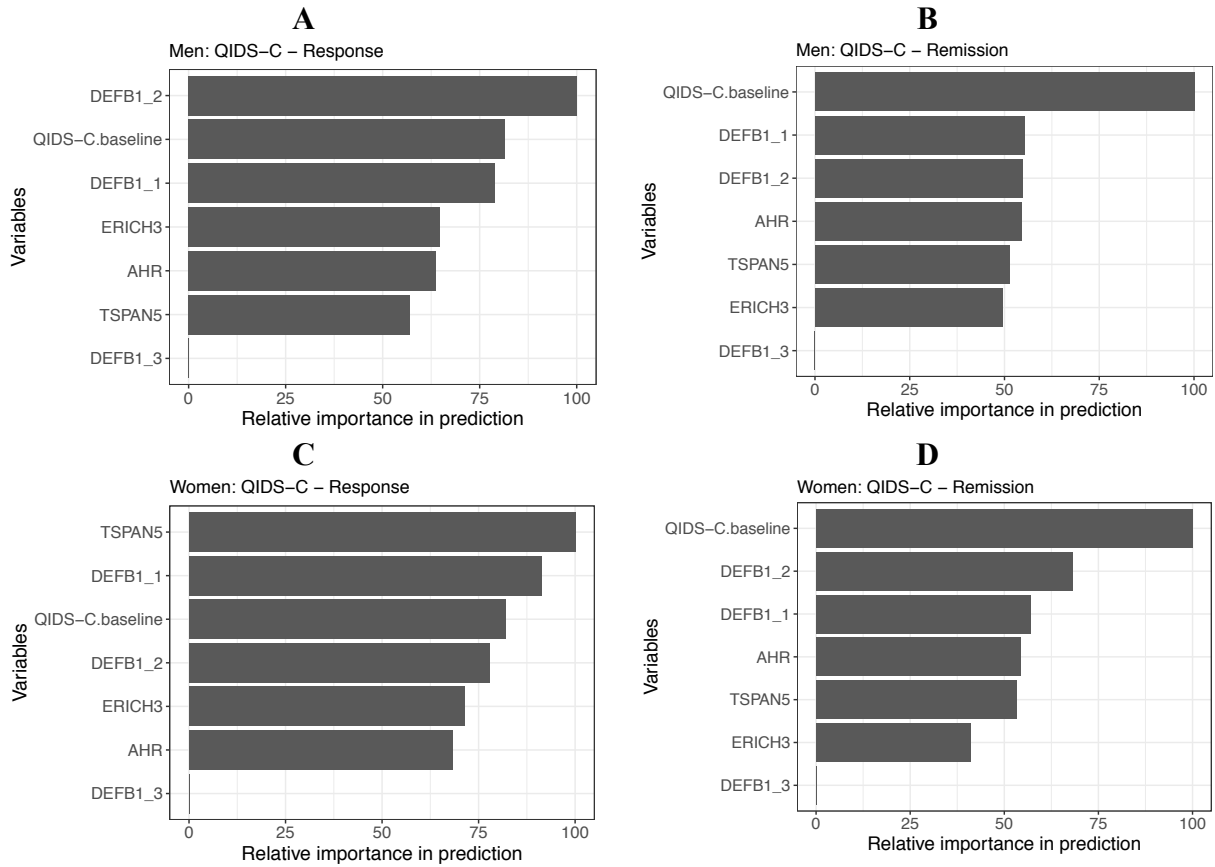
Fig. S4: Comparison of citalopram and escitalopram plasma drug concentrations between men and women with each depressive symptom severity cluster at 4 weeks (**Fig. A**) and 8 weeks (**Fig. B**).

rs numbers of SNPs: *AHR* (rs17137566), *TSPAN5* (rs10516436), *ERICH3* (rs696692), *DEFB1*_1 (rs5743467), *DEFB1*_2 (rs2741130) and *DEFB1*_3 (rs2702877)

Fig. S5: Importance of variables for predicting clinical outcomes measured using QIDS-C.